

A személyközi közelítés-távolítás azonosítása lokális nyelvtanok segítségével⁸¹

Pohárnok Melinda

PTE BTK Pszichológiai Intézet, Pécs, Ifjúság útja 6.
pomelin@freemail.hu

Kivonat: A narratív pszichológiai tartalomelemzés az élettörténeti narratívumok inherens jellegzetességének tekintheti a szereplők egymás viszonyában való mozgását. (Pohárnok, 2004)[1] Feltételezzük, hogy e mozgás két alapvető irányultsága – közelítés és távolítás – jól megragadható nyelvi struktúrákkal rendelkezik, amelyek a számítógépes nyelvi feldolgozás eszközeivel elérhetőek. Ugyanakkor ezen nyelvi elemek megfeleltethetőek a kapcsolati szabályozás és a self-szabályozás pszichológiai változójának. Az előadás bemutatja a NOOJ nyelvészeti fejlesztőkörnyezetben az MTA Nyelvtudományi Intézet munkatársainak segítségével folytatott munka jelenlegi fázisát, amelyben a személyközi közelítés-távolítás azonosítására szolgáló szemantikai szótárakat és ezekre épülő lokális nyelvtanokat hozunk létre.

1. Bevezetés

Az elbeszélést a narratív pszichológia szemszögéből az egyéni pszichikum szempontjából alapvető információkat hordozó entitásként és működési- és élményszervezésimódként tekintjük. Ugyanakkor azt is feltételezzük, hogy az elbeszélésben jól megragadhatóak azok a szövegelemek, amelyek valamely pszichológiai jelenségre utalnak. Míg a korábbi kvalitatív vizsgálatokban az elbeszélés pszichológiai jelentései pusztán az intuitív megértés számára voltak hozzáférhetőek, mi azt tűztük ki célul, hogy a számítógépes morfológiai-szintaktikai szövegelemzés, és – feldolgozás eszközével mérhetővé tegyük ezeket a pszichológiai jelentéseket. Korábbi vizsgálatok (pl. Pohárnok és mtsai., 2005)[2] rámutattak, hogy a tárgykapcsolati elbeszélések – kapcsolati epizódok – alapvető strukturális jellegzetességének és a narratív pszichológiai tartalomelemzés módszerével megragadható dimenziójának tekinthetjük a kapcsolati tér változásait. Ugyanakkor vizsgálataink azt is megerősítették, hogy a személyközi közelítés-távolítás az elbeszélésekben összefüggésbe hozható az elbeszélő személykapcsolati és érzelmi szabályozásának adekvátságával.

Munkánk alapvetése, hogy létezik egy olyan interperszonális vagy interaktív tér, amely mindig az én és a másik viszonya alapján szerveződik: a tér két végpontját az én és a másik adja meg, és egymás viszonyában való mozgásuk a kapcsolat alapvető sajátosságának tekinthető. Ez a tér, illetve az elbeszélő és a szereplők ebben való

⁸¹ A tanulmány az NKFP 6/074/2005 számú pályázat támogatásával készült.

mozgása kiemelkedik az elsősorban kapcsolati témájú élettörténeti narratívumokban, és így vizsgálhatóvá válik a gépi szövegfeldolgozás eszközeivel. A személyközi közelítés-távolítás az elbeszélésekben egyrészt konkrét, fizikai térben való mozgásokként jelenhet meg. A felé (vele) – tőle (nélküle) irányairól van szó. Másrészt megjelenhet az elbeszélő és a számára fontos mások közti érzelmi viszonyulások „irányában”, amely az én-elbeszélésekben az elbeszélő és a szereplők közti érzelmi viszony kifejezésében érhető tetten. A szeretet - gyűlölet ellentétes „irányai” adják ki ezt a dimenziót.

A kutatás jelenlegi fázisában a PTE Pszichológia Intézet és az MTA Nyelvtudományi Intézet együttműködésében folyik. Ennek a munkának egy szeletét mutatja be az előadás.

2. A nyelvi korpusz

A személyközi közelítés-távolítás azonosítására szolgáló nyelvtanok felépítését két irányból kezdtük el. Egyrészt egy olyan szövegtörzs felől, amelyben hangsúlyosan jelen vannak az én és a másik viszonyában zajló mozgások. Ebből a korpuszból még korábban kinyert minták alapján fogalmaztunk meg feltevéseket arra nézve, milyen szavak – szófajok -, szókapcsolatok felelhetnek meg a közelítésnek, illetve távolításnak. Másrészt a Magyar Nemzeti Szövegtár szófaji listáit használtuk fel, hogy azokat leszűkítve a közelítés-távolítás szempontjából releváns szótárakat építsünk.

Az a szövegtörzs, amelyben korábban manuális kódolással azonosítottuk a személyközi közelítésre-távolításra utaló kifejezéseket, egy 120 671 szóból álló élettörténeti epizódokat tartalmazó, élőnyelvi szöveg („BPDinterjúk”). Mindennek a kutatás szempontjából több lényeges vonatkozása van. Egyrészt a vizsgálati személyektől olyan epizódokat kértünk, amelyek életük egy meghatározó személyével kapcsolatosak (pl. „Mondja el első párkapcsolatának történetét!”), ez lehetővé tette, hogy az elbeszélő én és a szereplő másik viszonyára bőséges utalást találjunk a szövegekben. Ugyanakkor megnehezíti a személyközi mozgások referencialitásának felismerését azzal, hogy a közelítés-távolítás forrása és célja legtöbbször rejtve marad a szövegben, hisz a vizsgálati személynek feltett kérdés utal rá. Így azt a döntést hoztuk, hogy nemcsak a referencia személy közvetlen megjelölését fogadjuk el (pl. „az anyámat összecsókolgatom”), hanem közelítésnek tekintjük az E/1 személyű mozgást, amelynek célja E/3 vagy T/3 ragozott névmás, vagy határozószó, vagy névutó (pl. „*éjjel bementem és befeküdtem melléjük*”). Hasonlóan közelítésnek tekintjük az E/3 személyű mozgást, amelynek célja E/1 vagy T/1 ragozott névmás, vagy határozószó, vagy névutó (pl. „*jött, meg szaladt hozzám*”). Végül a korpusz sajátosságához tartozik, hogy élőnyelvi szöveg révén nyelvtanilag rosszul formált, töredékes, így ez sokszor akadálya volt annak, hogy egyes nyelvtanok pontosan alkalmazhatóak legyenek.

3. A személyközi közelítés-távolítás azonosítására szolgáló szemantikai szótárak

A kidolgozandó lokális nyelvtanoknak kétféle irányú mozgást kell azonosítaniuk, és egyaránt számításba kell venniük a fizikai mozgást (jön-megy), és az érzelmi “mozgást”(szeret-gyűlöl). Mindehhez az eddigiek során háromféle szófaji szótárt dolgoztunk ki.

Az Ige-szótárhoz a Magyar Nemzeti Szövegtár (MNSz) 10000 szavas ige listáját használtuk fel. A listát négy független kódoló tekintette át, és a következő kategóriák mentén értékelte a szavakat:

- 1.humán/nem-humán mozgás
- 2.iránnyal bíró/iránnyal nem bíró
- 3.közelítő/távolító
- 4.aktív/passzív.

	A	B	C	D	E	F	G
1075	leköt	2734	1	3			
1076	távozik	2729	1	2	5		
1077	társul	2724	1	2	4		
1078	megeszik	2718					
1079	megfogalmazód	2714					
1080	megtörlik/megtör	2713					
1081	megaláz	2701					
1082	elmenekül	2698	1	2	5		
1083	örököl	2697					
1084	beenged	2695	1	2	4		
1085	hozzáfűz	2694					
1086	ültet	2694	1	3			
1087	ment	2686					
1088	hever	2681	1	3			
1089	betilt	2671					
1090	elválik	2671	1	2	5		
1091	lebont	2670					
1092	fűződik	2654					
1093	letölt	2649					
1094	jósol	2646					
1095	méltat	2641					

1. ábra. Az igék szemantikai csoportokba sorolása

A kódolók közti egyeztetések után a nyelvtanokba beépíthető igék három csoportját tudtuk meghatározni:

1. Irány nélküli, vonzatfüggő humán cselekvés vagy állapot: *áll, vár, öltözik, marad, ül, leszalad, elköltözik*
2. Iránnyal bíró közelítő/közeledő mozgás: *érkezik, kísér, támaszt, beenged, csókol*
3. Iránnyal bíró távolodó/távolító mozgás: *kizár, elhagy, eltűnik, kirak, elereszt*

A következőkben az MNSz határozószó (n=10000) és a névutó (n=88) szótárait tekintettük át, és kigyűjtöttük, illetve a közelítésnek és távolításnak megfelelő csoportokba soroltuk a szavakat. Így a következő csoportok adódtak a határozószók esetében.

1. Közelítő határozószók – téri: *hozzá, neki, közelben, négyszemközt*
2. Közelítő határozószók – érzelmi: *aggodalmasan, biztatóan, féltőn, megnyugtatóan*
3. Távolító határozószók – téri: *egyedül, külön, messzebbre, odébb, távol*
4. Távolító határozószók – érzelmi: *elutasítóan, gunyorosan, kiábrándítóan, megcsalva.*

A névutók sok esetben átfedést mutattak a határozószókkal, így ott a szavaknak csak két kisebb csoportját tudtuk létrehozni.

1. Közelítő névutók: *alá, elé, elébe, felé, felől, köré, közé, között, mellé, mellett, mögé, mögött*
2. Távolító névutók: *aló, elől, innen, innét, kívül, közül, mellől, mögül, nélkül, túl*

4. A szótárokból felépített lokális nyelvtanok a NOOJ fejlesztőkörnyezetben

A kidolgozott szótárak segítségével az igecsoportok és a határozószók, illetve az igecsoportok és a névutók együtt járására terveztünk kidolgozni lokális nyelvtanokat. Az MTA Nyelvtudományi Intézete nyelvtechnológiai osztályának munkatársai által fejlesztett magyar nyelvű NOOJ fejlesztőkörnyezet morfológiai elemzője az igék és a névszók példányait meghatározott kimeneti jegyekkel látja el. A névszók esetén a ragnak megfelelő esetet, inflexiók toldalékokat – pl. többes szám, birtokos jelölve van - és képzőket használ. Az igék esetén többek között a szám, a személy, az igeidő és a mód morfológiai jegyek alapján meghatározott annotációja történik.

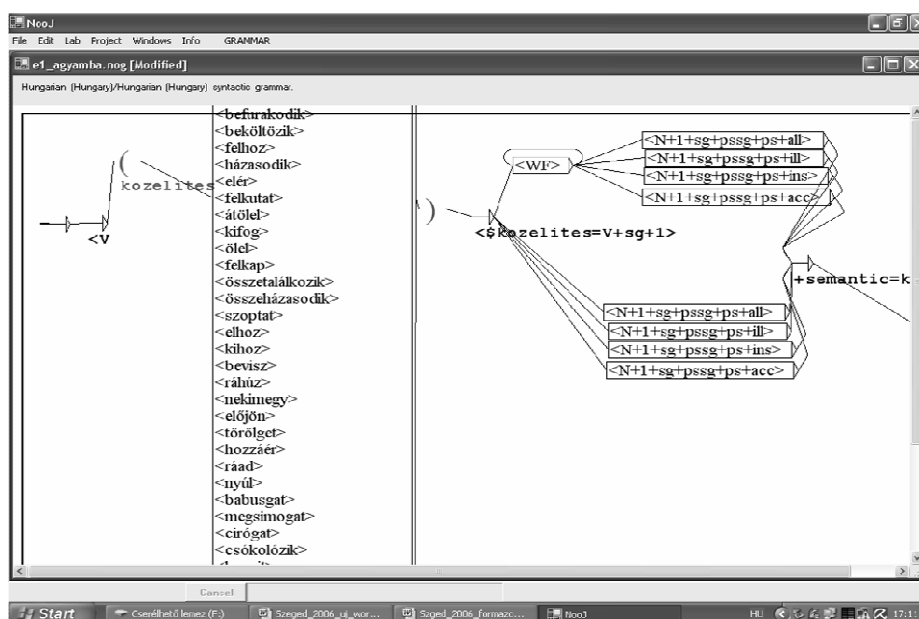
Mielőtt a szintaktikai nyelvtanok építésébe fogtunk, a Nyelvtudományi Intézet munkatársai rendelkezésünkre bocsátották az ún. „Jedlik-projekt” szövegtörzsének annotált szótárát. Ez tartalmazta azt a korpuszt is, amelyen a korábbi manuális kódolással elvégeztük már a személyközi közelítésre – távolításra utaló szóalakok, szókapcsolatok azonosítását. Ugyanakkor néhány szóalakot ebből a korpuszból sem ismert fel a NooJ. Például a meghatározó személyközi viszonyok szempontjából lényegesek közül az „anyámék”, „anyukámék”, „édesanyámék”, „barátnőmhöz” szóalakok voltak, amelyek annotációja a találati hibák elkerülése miatt nem történt meg. Módunk volt azonban ezeket a szóalakokat annotálni, majd beépíteni a korpusz szótárába.

Így végül lehetőségünk volt arra, hogy meghatározott számú, személyű igéket és meghatározott határozószókat, névutókat és névszók meghatározott esetű alakjait használjuk fel a nyelvtanok építéséhez.

4.1. Példák a személyközi közelítés azonosítására szolgáló lokális nyelvtanokból

Elsőként a közelítő igék azonosítására szolgáló egyszerű szintaktikai gráfot szerkesztettünk, azzal a céllal, hogy megtaláljuk azokat a közelítő igéket, amelyek szerepelnek a „BPDinterjúk”-ban. Így a kezdetben 596 szót tartalmazó teljes közelítő/közeledő-ige lista leszűkült 149 szóra, így könnyebben kezelhető vált a NooJ elemzője számára. Ezek után többek között a következő nyelvtanokba építettük be az igelistát:

1. Olyan szókapcsolat, ahol az E/1 vagy E/3 közelítő ige E/1 birtokos személyjellel ellátott, tárgyragos, vagy allativus, illativus, instrumentalis esetben lévő főnévvel áll együtt. Itt módosítani lehet, hogy a két szó közvetlenül egymás után, vagy bármely szóalakokat közrefogva van jelen a szövegben. Mivel a NooJ-ban az elemzés egységeként megadhatjuk a mondatot, ezért a mondat határai korlátozzák a keresett szóalakok közti szavak számát.



2. ábra. Ige-főnév közelítő gráf

Példák a gráf által azonosított szövegrészekből:

„és akkor *összetálcsoztam a barátommal* Miskolcon”

„oda *tartottam az arcomat*”

„soha nem *öleltem még meg anyámat*”

„hogy nem *fogom többet látni az apámat*”

Amennyiben a gráffal E/3 közelítő igéket E/1 birtokos személyjellel ellátott meghatározott esetben ragozott névszókkal együtt kerestünk, a következő példák adódnak:

„Aztán *beszaladt a hálószobámba* apám”

„*fogta a kezemet*, és akkor ez ilyen biztonságos érzés volt nekem”

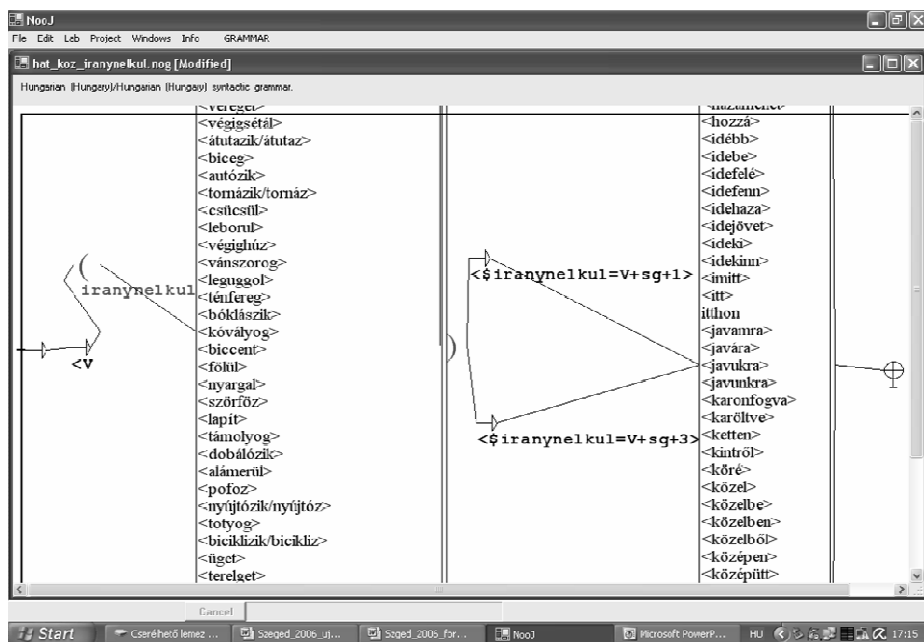
„hogy rendőrség *hozott mindig vissza apukámhoz*”

„de addig fiú nem *fogta meg a kezemet*”

A megtalált kifejezések egy része megfelel a fizikai közeledés szemantikai kritériumának, ugyanakkor a példákból is kitűnik, hogy meg kellett oldanunk a tagadás kezelését. Erre egyelőre egy önálló negáció gráfot tudunk használni, amely megtalálja egy másik gráf által azonosított szövegrészek egyszerű tagadó formáit. A példák esetében kiemelni a „de addig fiú *nem fogta meg a kezemet*” és „*hogy nem fogom többet látni az apámat*” eseteket. Így ezek más szemantikai kategóriába kerülhetnek. A tagadás kezelése a személyközi közelítés-távolítás esetében azért fontos, mert a közelítés tagadása távolítást jelenthet - ahogy az előbbi példákból is kitűnhet.

A megtalált példák arra is felhívják a figyelmet, hogy a szavak azonos alakúság a beleszólhat az eredményekbe: a „fog”-ra nemcsak mint aktív igére, hanem a jövő időt kifejező segédigére is találatot kapunk.

2. Olyan szókapcsolat, ahol a közelítő, vagy irány nélküli humán cselekvést, állapotot kifejező E/1 és E/3 ige közelítő határozószóval áll együtt. A határozószó definíciójának bizonytalansága miatt ide kerültek azok az esetek is, amikor az ige után ragozott névmás áll (pl. *hozzám, vele*), mert ezeket határozószóként és névmásként is annotálták a szótárban.



3. ábra. Ige-határozószó közelítő gráf

Példák a gráf által azonosított szövegrészekből:

„Aztán meg este *bejött hozzám*”
 „Tehát így *lefeküdtem mellé* a földre”
 „és akkor így *vártam rá*, hogy elvisz”
 „Aznap mikor *jöttem hazafelé*”

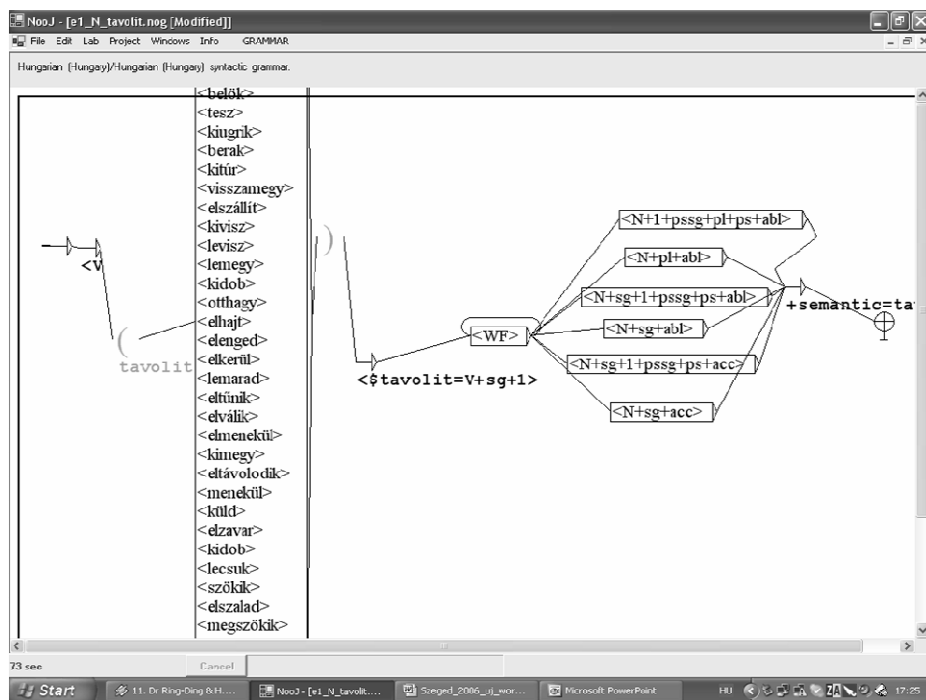
„Akkor *álltam neki* életemben először főzni”
 „és nagyon sokáig mérges is *volt rám*”
 „hogy ha az életben még *kapok rá* lehetőség”
 „Ő tudta, hogy *van neki* valakije”

Ahogy az idézett példák második fele is mutatja, ebbe a nyelvtanba további megszorításokat kell beépítenünk. Szükségesnek tűnik úgy megszűrni a határozószók, ragozott névmások szótárát, hogy az állandósult szókapcsolatokat és az elváló igekötőként is funkcionáló névmások egy részét (pl. *nekiáll, rákap, rájön*) kiemeljük. Ugyanakkor további megszorításokat lehet elérni azzal, hogy ha kiemeljük a ragozott névmásokat. Ezeknek a számát, személyét később megadva ugyanis ki fogjuk tudni zárni azokat az eseteket, amelyekben azonos számú és személyű az ige és a cél is (pl. hogy ő bármit *adhat neki*), így nem az elbeszélő én és a másik viszonyára utalnak.

4.2. Példák a személyközi távolítás azonosítására alkalmas lokális nyelvtanokból

A távolítás azonosításához az előzőekhez hasonlóan a távolító illetve irány nélküli humán cselekvést vagy állapotot kifejező igéket, a távolító határozószókat és a ragozott névmások meghatározott csoportját használtuk fel. Így többek között a következő lokális nyelvtanok adódnak:

1. Olyan szókapcsolat, ahol az E/1 vagy E/3 távollétást kifejező ige E/1 birtokos személyjellel ellátott, ablativus, delativus vagy accusativus esetben lévő főnévvel áll együtt. Itt változtattuk a névszó és az ige egymásra következőségének sorrendjét, és köztük lévő szóalakok előfordulási lehetőségét is beiktattuk.



4. ábra. Ige-névszó távolító gráf

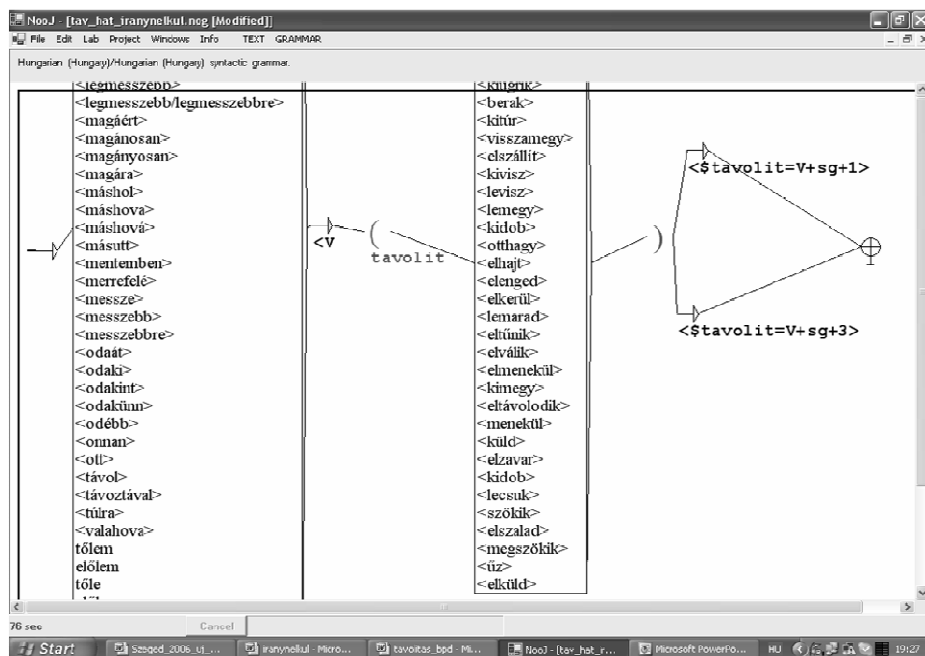
Az első lépésben itt is leszűkítettük az általános távolítás szótárt a korpuszban megtalálható szavakra az egyszerűbb kezelhetőség érdekében. Ennek következtében azonban olyan szemantikailag kevésbé differenciált igék maradtak csak benne a korpuszra illesztett szótárban, amelyek miatt sok érvénytelen találatot kaptunk a helyes találatok mellett.

„De akkor is én *hagytam ott a fiút*”
 „Ezért *megyek egyik pasitól* a másik után”
 „ha *elválok a férjemtől*”
 „amikor már külön *mentem a férjemtől*”
 „hogya *mamámtól elkerültem*”

 „nagyon sok *szenvedéstől menekültem* így meg.”
 „később francia *bokszt üttem*”
 „hogya mindennap *elementem venni egy csomó édességet*”
 „de égve *hagytam a villanyt*”

A hibás találatok kizárása azzal oldható meg, hogy a nyelvtan névszói részét tovább szűkítjük azokra a főnevekre, amelyek jelentős másokra – szülők, családtagok, szerelmi partner, orvosok - vonatkoznak. Ezt a NooJ-ban két módon tehetjük meg. Egyrészt a korpuszban található szóalakokat egyenként visszük be a gráfba, vagy létrehozuk a „jelentős másik” szemantikai kategóriát úgy, hogy ennek megfelelő annotációval látjuk el a korpusz szótárában az ehhez a kategóriához tartozó szóalakokat. A kutatás következő fázisában tervezünk létrehozni a fenti névszói kategóriákra alkalmazható szótárt, amely alapján új szemantikai annotációkkal és új szóalakokkal bővíthetjük a jelenleg használt NooJ szótárt.

2. Olyan szókapcsolat, ahol a távolító, vagy irány nélküli humán cselekvést, állapotot kifejező E/1 és E/3 ige távolító határozószóval áll együtt. A határozószó definíciójának bizonytalansága miatt ide kerültek azok az esetek is, amikor az ige után ragozott névmás áll (pl. *tőle, magától, előlem*). Az ige és a névszó sorrendjének itt is kétféle változatát használtuk. Az erre a szókapcsolatra kidolgozott gráfok egyik változatát mutatja az ábra.



5. ábra. Határozószó-ige távolító gráf

Példák a gráf által azonosított szövegrészekből:

„úgy éreztem biztos, hogy *magamra maradtam*”

„És akkor *megszöktem onnan*”

„hogy éjjel is *lökött magától*”

„a lányom, az már talán *eltávolodott tőlem*”

„Anyukám mindig sietett, *futott valahova*”

„hogy terhes *vagyok tőle*”

„Mert *volt külön szobám*”

„Hát sok jó élményem nem *maradt belőle*”

„és akkor utána *megy magától*”

„pár hónap alatt *lefut belőlem*”

A találatok arra hívják fel a figyelmet, hogy az irány nélküli humán cselekvésekhez sorolt létigék alkalmazása miatt adódik a legtöbb hibázás, így érdemes lenne a jövőben a létigékre és a hozzájuk kapcsolódó mind közelítést, mind távolítást kifejező szóalakokra külön algoritmusokat építeni. Ugyanakkor a távolítás kapcsán – ahogy korábban már említettük – különösen fontos lesz a tagadás rugalmas használata. Ennek kidolgozása is a további munka része.

A továbbiakban ezeknek a nyelvtanoknak a finomítását tervezzük, többek között a tagadás beépítésével, a sok téves riasztás adó igealakok – például a létigék – alkalmas előfordulásainak kiszűrésével és a már említett főnévi szemantikai kategóriák felhasználásával. Végül, mivel a jelenlegi nyelvtanok a fizikai mozgások azo-

nosítására alkalmasak, dolgozunk az érzelmi közelítésre-távolításra utaló szavak szótárának felépítésén, és szintaktikai nyelvtanokba való beépítésén.

Bibliográfia

1. Pohárnok M.: A térben való mozgás narratív dimenziójának vizsgálata borderline és depressziós betegek élettörténeti epizódjaiban. IN: Erős F. (szerk.): Az elbeszélés az élmények kulturális és klinikai elemzésében. Akadémiai Kiadó, Bp., (2004) 153-167
2. Pohárnok M., Nagy L., Bóna A., Naszódi M., Kis B., László J.: A kapcsolati mozgások számítógépes nyelvészeti vizsgálata élettörténeti narratívumokban. A LAS-vertikum közelítés-távolítás modulja. Pszichológia, (2005) 2:157-171.